

Editor

César Antônio Pereira

Conflito de interesses

A autora declara não haver conflito de interesses.

Recebido

29 jan. 2023

Versão final

22 mar. 2023

Aprovado

2 abr. 2023

Recuperação de informação na Ciência da Informação: produção acadêmico-científica brasileira (2012-2021)

Information retrieval in Information Science: Brazilian academic-scientific production (2012-2021)

Patrícia Nascimento Silva¹ 

¹ Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Programa de Pós-Graduação em Gestão & Organização do Conhecimento. Belo Horizonte, MG, Brasil. E-mail: patricians@ufmg.br

Como citar este artigo/How to cite this article: Nascimento Silva, P. Recuperação de Informação na Ciência da Informação: produção acadêmico-científica brasileira (2012-2021). *Transinformação*, v. 35, e237336, 2023. <https://doi.org/10.1590/2318-0889202335e237336>

Resumo

A recuperação de informação é uma área de estudo nativa da ciência da computação, cujo termo foi criado somente em 1951. No entanto, por mais de 5 mil anos, a humanidade vem organizando a informação para sua busca e recuperação. Com o advento dos computadores, o termo que era de interesse apenas de bibliotecários teve seu desenvolvimento acelerado. Desafios atuais envolvendo novos dispositivos e fenômenos, como o *Big Data*, continuam incitando estudos na área para promover aos usuários o acesso às informações. Este artigo tem o objetivo de identificar a produção acadêmico-científica brasileira sobre a recuperação de informação na ciência da informação nos últimos 10 anos. A Pesquisa é descritiva e exploratória, com abordagem quali-quantitativa, por meio de uma revisão bibliográfica integrativa na Biblioteca Digital Brasileira de Teses e Dissertações no período de 2012 a 2021. Foram selecionados 69 documentos, categorizados em 27 temáticas. A maior frequência de documentos foi observada para as seguintes temáticas: estudo de usuários, mecanismos de busca, representação da informação, modelo de recuperação e representação temática. Observou-se que as temáticas mais recentes são apresentadas a partir de 2016, e que o maior número de trabalhos está vinculado à Universidade Federal de Minas Gerais. No contexto da Ciência da Informação, a recuperação de informação tem forte ligação com a organização e a representação da informação e compõem uma relação quase que indissociável. O estudo é um aporte para novos pesquisadores, contribuindo para publicizar as contribuições existentes e instigando novas reflexões sobre a recuperação de Informação na Ciência da Informação.

Palavras-chave: Ciência da Informação. Recuperação de Informação. Revisão de Literatura.

Abstract

Information retrieval is an area of study native to computer science, whose term was coined only in 1951. However, for over 5000 years, humanity has been organizing information for its search and retrieval. With the advent of computers, the term that was only of interest to librarians had its

development accelerated. Current challenges involving new devices and phenomena, such as Big Data, continue to encourage studies in the area to provide users with access to information. This paper aims to identify the Brazilian academic-scientific production on information retrieval in information science in the last 10 years. The research is descriptive and exploratory, with a quali-quantitative approach, through an integrative bibliographic review in the Brazilian Digital Library of Theses and Dissertations from 2012 to 2021. Sixty-nine documents were selected, categorized into 27 themes. The highest frequency of documents was observed for the themes: user studies, search engines, information representation, retrieval model, and thematic representation. It was observed that the most recent themes are presented as of 2016 and the largest number of works is linked to the Federal University of Minas Gerais. In the context of Information Science, information retrieval has a strong connection with the organization and representation of information and makes up an almost inseparable relationship. The study is a resource for new researchers, helping to publicize existing contributions and instigating new reflections on the retrieval of Information in Information Science.

Keywords: Information Retrieval. Information Science. Literature review.

Introdução

A recuperação de informação é uma área ampla, nativa da ciência da computação, que se concentra em promover aos usuários acesso às informações de seu interesse (Baeza-Yates; Ribeiro-Neto, 2013). Para Saracevic (1999), a recuperação de informação pode ser considerada uma vertente tecnológica da ciência da informação, resultante da sua relação com a ciência da computação.

A Ciência da Informação é considerada uma área multidisciplinar que recebe contribuições de diferentes disciplinas e campos científicos. Segundo Araújo (2003), a área é tradicionalmente classificada na ciência social aplicada, perpassando questões interdisciplinares e aproximando-se da ciência pós-moderna. Com a era digital e a crescente apropriação da informação pelos usuários, a recuperação de informação recebeu contribuições relevantes tanto da ciência da computação quanto da ciência da informação de forma que essa relação de simbiose entre as duas áreas, inicialmente conflituosa, é cada vez mais marcante e indissociável, sobretudo a partir dos anos 1990.

Com o surgimento da *Web* e a evolução propiciada pela tecnologia, novos recursos foram disponibilizados e incorporados aos sistemas de recuperação de informação, ampliando as formas de acesso, em volumes cada vez mais expressivos, distribuídos e complexos, no contexto cognominado *Big Data*. Todavia, para a recuperação de informação ser considerada uma tarefa natural e simples, na atual era da informação, foi imprescindível incorporar aos sistemas procedimentos clássicos da biblioteconomia, da documentação e da recuperação de informação (Ferneda, 2012). Para tanto, lidar com a representação e a organização da informação, em diferentes contextos e formatos, é essencial para garantir relevância e confiabilidade na recuperação de informação, além de atender às necessidades informacionais de um determinado público.

A *Web* semântica, proposta em 2001 por Tim Berners-Lee, James Hendler e Ora Lassila (Berners-Lee; Hendler; Lassila, 2001), trouxe proposições para estrutura da *Web* atribuindo significado passível de interpretação por pessoas e máquinas. Com essa proposta, foi criado um conjunto de conceitos estruturados em camadas, o que ampliou as possibilidades de recuperação de informação e a complexidade envolvida nesta atividade.

A temática e a área de pesquisa recuperação de informação abarcam desafios atuais da ciência da informação. Investigar a sua evolução é importante para detectar o progresso da ciência da informação, visualizar novas associações e desenvolver novos recursos. Dessa forma, o objetivo desta pesquisa foi identificar a produção acadêmico-científica brasileira

sobre a recuperação de informação, na ciência da informação, nos últimos 10 anos, remetendo à seguinte questão investigativa: Como a recuperação de informação tem sido abordada na produção acadêmico-científica brasileira da ciência da informação entre 2012 e 2021? Buscou-se, especificamente, (1) investigar as temáticas relacionadas à área e (2) identificar as características dos estudos produzidos no Brasil. Este estudo justifica-se para mapear as pesquisas mais recentes e elucidar a evolução da área no Brasil e pretende contribuir como aporte a novas reflexões sobre a recuperação de informação na Ciência da Informação.

Recuperação de Informação

A recuperação de informação é uma área de conhecimento recente e em constante evolução. O termo “*Information Retrieval*”, recuperação de informação, foi criado em 1951 por Calvin Mooers. Foi nesta década que vários cientistas, engenheiros e bibliotecários se empenhavam na busca de soluções para problemas enfrentados por Otlet, no início de século, e atualizados por Bush após a Segunda Guerra Mundial (Ferneda, 2012). No entanto, é importante destacar que, por mais de 5000 anos, a humanidade vem organizando a informação para sua busca e recuperação. A indexação de tabuletas de argila, hieróglifos, rolos de papiros e livros, em bibliotecas milenares, já evidenciava as primeiras práticas de recuperação de informação (Baeza-Yates; Ribeiro-Neto, 2013).

Com o advento dos computadores, o termo que era de interesse apenas de bibliotecários e especialistas de informação teve seu desenvolvimento acelerado. O termo se tornou uma área abrangente da ciência da computação que se concentrou em prover aos usuários acesso fácil às informações de seu interesse (Baeza-Yates; Ribeiro-Neto, 2013):

A Recuperação de Informação trata da representação, armazenamento, organização e acesso a itens de informação, como documentos, páginas *Web*, catálogos online, registros estruturados e semiestruturados, objetos multimídias, etc. A representação e a organização dos itens de informação devem fornecer aos usuários facilidade de acesso às informações de seu interesse (Baeza-Yates; Ribeiro-Neto, 2013, p.1).

Todavia, a própria definição da área indica sua amplitude e simbiose com a Ciência da Informação. Segundo Ferneda (2003), a Recuperação da Informação pode se constituir como um importante domínio interdisciplinar entre a Ciência da Computação e a Ciência da Informação, a partir dos sistemas quantitativos de recuperação da informação, desde que, além dos sistemas e modelos de recuperação quantitativos, sejam discutidos os modelos dinâmicos de recuperação da informação, que podem considerar as formas significativas do objeto informacional e os processos e fluxos que os envolvem.

De acordo com Saracevic (1996), esta disciplina se aproxima da Ciência da Informação a partir dos processos de recuperação da informação no uso dos computadores e automatização das coisas, os quais, por sua vez, se constituem como elemento tecnológico diretamente relacionado aos sistemas de informação. Nesse sentido, Wersig e Neveling (1975) afirmam que a Ciência da Computação se aproxima da Ciência da Informação, a partir da relação do papel que a tecnologia tem na gestão da informação.

Com isso, desde o surgimento do termo, na década de 1950, percebe-se a evolução da área de recuperação de informação. Os sistemas de recuperação de informação, representados por “conjuntos de operações consecutivas executadas para localizar, dentro da totalidade de informações disponíveis aquelas realmente relevantes” (Cesarino, 1985, p.1), acompanharam o surgimento de novas tecnologias e a evolução dos sistemas de informação, à medida que a necessidade do usuário e os problemas de recuperação da informação se modificam. A *Web* foi um

dos marcos da evolução da área, com o surgimento de novos desafios, marcados, inicialmente, pelos mecanismos de busca em páginas e novas estruturas de representação a partir da *Web Semântica* (Ferneda, 2012).

Nesse ínterim, nos últimos 10 anos, período compreendido entre 2012 e 2021, observam-se diversos desafios na recuperação de informação relacionados ao grande volume de informações, aos novos recursos e dispositivos tecnológicos e ao uso acentuado de técnicas e algoritmos computacionais (inteligência artificial, processamento de linguagem natural, aprendizado máquina), o que reitera a necessidade de constante evolução da área e motivou este estudo, que traça um panorama da produção acadêmico-científico, no âmbito da Ciência da Informação, no Brasil.

Procedimentos Metodológicos

Com objetivo de investigar a produção acadêmico-científica brasileira, no período de 2012 a 2021, optou-se por realizar uma pesquisa descritiva e exploratória, com abordagem quali-quantitativa, por meio de uma revisão bibliográfica integrativa. Esse tipo de pesquisa bibliográfica, oriundo da prática baseada em evidências, apresenta os estudos científicos atuais, para um determinado domínio, a fim de compreendê-los e contribuir para o desenvolvimento de novas pesquisas e práticas na área, conforme os resultados já obtidos em outros estudos. Além disso, permite a inclusão de diferentes tipologias de pesquisa (experimentais, não-experimentais, teóricas, empíricas) com o propósito de compreender o fenômeno de maneira completa (Whittemore; Knafl, 2005), permitindo que o profissional de determinado domínio se mantenha atualizado quanto aos conhecimentos científicos produzidos em sua área (Oliveira; Cavalcante, 2022).

A revisão integrativa deste estudo contemplou as teses e as dissertações indexadas na Biblioteca Digital de Teses e Dissertações (BDTD), com uma abordagem qualitativa e quantitativa. A BDTD é uma base de dados que integra os sistemas de informações de teses e dissertações das Instituições de Ensino e Pesquisa no Brasil e possibilita à comunidade brasileira de ciência e tecnologia a publicação das teses e dissertações produzidas no país e no exterior, proporcionando visibilidade a produção científica nacional (Instituto Brasileiro de Informação em Ciência e Tecnologia, 2023). Dessa forma, a base pode ser considerada a fonte mais completa sobre a produção acadêmico-científica no Brasil. Baseada nessa representatividade e na natureza da fonte de informação, que indexa todas as dissertações e teses defendidas nos programas de pós-graduação brasileiros, a BDTD foi selecionada como fonte dessa revisão integrativa. Outros meios de publicação, como anais de eventos e periódicos científicos, também são fontes de informação da produção nacional, mas não foram objetos deste estudo, que pesquisou os seguintes trabalhos acadêmicos: teses e dissertações.

Os autores Mendes, Silveira e Galvão (2008) definem o método de revisão integrativa como o mais amplo de revisão e especificam seis etapas para realizá-lo: Identificação do tema e seleção da hipótese ou questão de pesquisa; Estabelecimento de critérios para inclusão e exclusão de estudos; Definição das informações a serem extraídas dos estudos selecionados; Avaliação dos estudos selecionados; Interpretação dos resultados e Apresentação da revisão em uma síntese do conhecimento. O Quadro 1 apresenta as etapas e a descrição dos critérios adotados nesta revisão integrativa.

Observando os recursos da busca avançada da BDTD, foi identificado que a pesquisa poderia utilizar os campos a seguir: título, autor, assunto, resumo em português, resumo em inglês, editor, ano de defesa. Dessa forma, a busca utilizou os campos título, resumo em português, resumo

Quadro 1 – Definições da Revisão Integrativa.

Definições da Revisão Integrativa	
Etapas	Descrição
Identificação do tema e seleção da hipótese ou questão de pesquisa	Delinear um panorama da produção acadêmico-científica brasileira sobre a recuperação de informação e seus elementos constitutivos mais significativos, publicados no âmbito da Ciência da Informação entre 2012 e 2021.
	Questão norteadora: Como a recuperação de informação tem sido abordada na produção acadêmico-científica brasileira da ciência da informação nos último 10 anos?
Estabelecimento de critérios para inclusão e exclusão de estudos/ amostragem ou busca na literatura	Base de dados: BDTD
	Período: 2012 a 2021
	Idioma: Português, Inglês, Espanhol
	Inclusão: Documentos sobre Recuperação de Informação e Ciência da Informação. Exclusão: Documentos de outras áreas do conhecimento ou documentos que fizeram menção aos conceitos, mas que não tratavam da temática.
Definição das informações a serem extraídas dos estudos selecionados	Atributos selecionados: título, resumo, palavras-chave
	Expressões de busca: ("recuperação de informação" and "ciência da informação") OR ("recuperação da informação" and "ciência da informação") OR ("information retrieval" AND "information science") OR ("recuperación de información" and "ciencia de la información")
	Leitura dos títulos e resumos dos documentos recuperados com o intuito de verificar a pertinência do objetivo geral do trabalho com a questão de pesquisa deste estudo.
Avaliação dos estudos incluídos na revisão integrativa	Categorização dos estudos conforme a principal temática abordada e leitura dirigida sobre a estrutura dos trabalhos (capítulos e demais temáticas abordadas) e resultados obtidos.
Interpretação dos Resultados	Discussão dos principais conceitos sobre a recuperação de informação e seus elementos constitutivos mais significativos e resultados da pesquisa.
Apresentação da revisão/síntese do conhecimento	Apresentação dos principais resultados evidenciados da análise dos trabalhos selecionados.

Fonte: Elaborado pela autora. Belo Horizonte (2023).

em inglês, assunto e ano de defesa. Como não existia a possibilidade de realizar a busca por palavras-chave, foi utilizado o campo assunto, no idioma português, predominante da base. As expressões de busca combinaram os termos recuperação de informação e Ciência da Informação e suas variações no período de 2012 a 2021. O Quadro 2 apresenta o resumo das buscas realizadas na BDTD, em janeiro de 2023.

Optou-se por fazer a busca com uma janela temporal de dez anos, por ser um período representativo que permite observar mudanças e movimentos incorporados aos estudos publicados. Considerando que a busca foi realizada em janeiro de 2023 e que os trabalhos defendidos em 2022 poderiam ainda não estar disponibilizados na BDTD, foi delineado o período de 2012 a 2021 que compreenderia os últimos 10 anos, cumprindo com o objetivo de trazer os estudos mais recentes da área.

Uma busca manual complementar também foi realizada no Google Acadêmico, para conferir os resultados identificados na BDTD. Foi utilizada a expressão "recuperação" "informação" no período de 2012 a 2022 que retornou 99 mil resultados, para todos os tipos de documentos. As 10 primeiras páginas indexadas foram conferidas e somente 6 trabalhos, do tipo dissertação ou tese, não foram identificados na busca realizada na BDTD. No entanto, foi possível identificar que

os trabalhos não foram recuperados na BDTD, pois não tinham o termo Ciência da Informação no título, resumo ou assunto, indicando que não havia problemas na recuperação dos documentos na BDTD. É importante destacar que essa conferência manual e complementar não interferiu nos resultados coletados na BDTD.

Conforme destacado por Paes e Tabosa (2015), há deficiências na recuperação de informações na BDTD/UFC, por exemplo, e as causas podem estar associadas a diversos fatores que vão desde a arquitetura definida na concepção da base, perpassando os metadados utilizados na representação dos documentos, até o trabalho dos indexadores que alimentam a base de dados. No entanto, essas questões não são objeto desta pesquisa, que considerou o instrumento BDTD como uma fonte de dados fidedigna. Contudo, é de conhecimento dos autores que pode haver limitações, devido a algum atraso na disponibilização dos trabalhos que foram defendidos em 2021, mas ainda não foram indexados na BDTD, ou erros na alimentação da base de dados.

Cada expressão de busca com os documentos recuperados gerou um arquivo de exportação. Todos os documentos exportados foram agrupados em uma única planilha, por meio da ferramenta Knime. Todos os arquivos manipulados foram disponibilizados no repositório criado para esta pesquisa, disponível no endereço: <https://github.com/pprof2022/RI_2012-2021.git>, conforme nomenclatura indicada no campo "ID" do Quadro 2.

Quadro 2 – Busca realizadas na BDTD.

ID	Expressão	Atributo	Documentos Recuperados	Observação
R1	(Título:"recuperação de informação" AND "ciência da informação")	Título	1	
R2	(Título:"recuperação da informação" AND "ciência da informação")	Título	1	
-	(Título:"information retrieval" AND "information science")	Título	0	
-	(Título:"recuperación de información" AND "ciencia de la información")	Título	0	Filtro: Ano defesa (2012 a 2021)
R3	(Resumo Português:"recuperação de informação" AND "ciência da informação")	Resumo português	229	
R4	(Resumo inglês:"information retrieval" AND "information science")	Resumo inglês	78	
R5	(Assunto:"recuperação de informação" AND "ciência da informação")	Assunto	95	
R6	(Assunto:"recuperação da informação" AND "ciência da informação")	Assunto	95	
Documentos duplicados na planilha (246)				
R7	Síntese das buscas na BDTD	Documentos duplicados identificados durante a leitura (18)		
			Total = 235	

Fonte: Elaborado pela autora. Belo Horizonte (2023).

Consoante os critérios definidos para essa revisão integrativa, inicialmente, a análise consistiu na leitura do título e do resumo dos 253 trabalhos recuperados, observando o objetivo geral do trabalho e marcando os que estavam no escopo da revisão. Durante a leitura dos títulos, foram identificados 18 trabalhos duplicados, pois alguns documentos estavam cadastrados em duas instituições de Ensino. Dessa forma, a amostra foi composta por 235 trabalhos. Dentre esses, foram selecionados 69 trabalhos, cujo objetivo geral, informado no resumo do trabalho e no título, estavam relacionados à recuperação de informação.

Em um segundo momento, os documentos relevantes foram classificados, conforme a principal temática abordada. Destaca-se que o nome da temática foi proposto pela autora, baseado em seu conhecimento prévio sobre os assuntos mais recorrentes na área de Ciência da Informação. Por fim, foi realizado o *download* dos 69 trabalhos selecionados e uma leitura dirigida sobre a estrutura dos trabalhos (capítulos e demais temáticas abordadas), além dos resultados obtidos nas pesquisas. A leitura também apontou alguns ajustes sobre as temáticas identificadas inicialmente para cada trabalho, que foram adequadas para serem apresentadas nos resultados (Quadro 3, 4 e Figura 1).

Quadro 3 – Produção acadêmico-científica analisada.

1 de 3

ID	Título	Classificação (Temáticas)	Ano de defesa
D1	Batendo um papo com a informação: o uso dos <i>chatbots</i> para a recuperação da informação e a contribuição da Ciência da Informação nesse processo.	<i>Chatbot</i>	2020
D2	Preservação e recuperação de informação em fontes de informações digitais: estudo de caso do <i>Greenstone</i> .	Repositórios digitais	2012
D3	A representação e a recuperação da informação: bases, diálogos e contribuições para o fazer arquivístico.	Arquivologia	2019
D4	Recuperação de informação de música e dados ID3: possíveis aplicações.	Representação descritiva	2015
D6	A representação temática da informação em TCC do Curso de Direito da UFPB: análise do vocabulário controlado no Repositório Eletrônico Institucional da UFPB.	Repositórios digitais	2013
D7	Um modelo híbrido de recuperação de informação e conhecimento baseado na síntese de mapas conceituais obtidos por operações de transformação de redes complexas orientadas por busca de relacionamentos entre termos de consulta em bases de dados ligados.	Modelo de recuperação	2016
D8	O campo da Ciência da Informação: contribuições, desafios e perspectivas da mineração de dados para o conhecimento pós-moderno.	Epistemologia	2015
D10	Representação e recuperação da informação em centros de memória.	Representação da informação	2015
D11	Recuperação de informação em jornais on-line: percepção sobre atributos de pesquisa em mecanismos de busca.	Mecanismos de busca	2015
D13	Extroversão e descoberta: visualização aplicada em sistemas de recuperação da informação.	Visualização da informação	2019
D14	Representação e recuperação de informação em acervos digitais nos contextos da <i>web</i> semântica e <i>Web</i> pragmática: um estudo crítico	Repositórios digitais	2015
D15	Interface multimodal de interação humano-computador em sistema de recuperação de informação baseado em voz e texto em português.	Interação humano-computador	2013
D17	Impacto do fator emocional quando da recuperação de informação da <i>home page</i> do Departamento de Doenças Sexualmente Transmissíveis, AIDS e Hepatites Virais da Secretaria de Vigilância em Saúde do Ministério da Saúde.	Estudo de usuários	2016
D19	A relevância da política de indexação para representação e recuperação da informação: uma revisão sistemática	Indexação	2019
D20	Influência dos fatores cognitivos e afetivos no processo de recuperação da informação: um estudo de caso em instituições de ensino superior na cidade de Chimoio, Manica, Moçambique.	Estudo de usuários	2015
D21	Organização e Recuperação da Informação Musical: o <i>incipit</i> como elemento de representação.	Representação da informação	2018
D23	Avaliação da arquitetura da informação de bibliotecas digitais de teses e dissertações: o caso da BDTD do IBICT.	Arquitetura da informação	2014
D24	Sistema de recomendação de objetos de aprendizagem digitais para <i>e-learning</i> : um estudo de caso em curso superior à distância da UFSC.	Sistemas de recomendação	2017
D25	Lexicalização de ontologias: o relacionamento entre conteúdo e significado no contexto da recuperação da informação.	Ontologias	2015
D26	Avaliação da usabilidade na interação e recuperação da informação dos usuários pós-graduandos no Repositório Institucional da Universidade de Brasília.	Estudo de usuários	2016

Quadro 3 – Produção acadêmico-científica analisada.

ID	Título	Classificação (Temáticas)	Ano de defesa
D30	Avaliação da indexação a partir dos elementos de exaustividade e especificidade pela revocação e precisão na recuperação da informação: sob a ótica dos usuários no sistema de bibliotecas da UFPA.	Indexação	2019
D34	Recuperação da informação: estudo da usabilidade na base de dados <i>Public Medical</i> (PUBMED).	Estudo de usuários	2014
D36	A caracterização dos objetos virtuais de aprendizado como recurso informacional: um estudo sobre a recuperação da informação na educação à distância.	Representação descritiva	2017
D38	Representação da informação musical: subsídios para recuperação da informação em registros sonoros e partituras no contexto educacional e de pesquisa.	Representação descritiva	2012
D39	Atuação bibliotecária no tratamento temático da informação em unidades informacionais: um estudo comparativo qualitativo-quantitativo.	Representação temática	2013
D41	Interoperabilidade e mapeamentos entre sistemas de organização do conhecimento na busca e recuperação de informações em saúde: estudo de caso em ortopedia e traumatologia.	Mecanismos de busca	2015
D42	Os Nativos digitais e a recuperação da informação científica on-line.	Estudo de usuários	2012
D43	Organização do conhecimento em bibliotecas digitais de teses e dissertações: uma abordagem baseada na classificação facetada e taxonomias dinâmicas.	Representação temática	2013
D46	Formas de redução léxica na redação de resumos de alta em prontuário do paciente visando à indexação e recuperação da informação.	Representação da informação	2016
D47	Um estudo sobre a interação dos usuários com a interface e o sistema de busca do Portal de Periódicos da Capes.	Mecanismos de busca	2014
D49	Estratégias de busca por assunto no Sistema Único de Informação: estudo com os usuários internos da Procuradoria da República nos Estados da Paraíba, Rio Grande do Norte e Pernambuco.	Estudo de usuários	2019
D51	Projeto de sistemas de recuperação de informação corporativa: uma abordagem de análise de domínio baseada na análise facetada.	Representação temática	2014
D54	Memória, mudança linguística versus recuperação em documentos de arquivo no longo prazo.	Linguística	2015
D57	Uso da linguagem documentária na busca da informação em bibliotecas universitárias: a perspectiva dos deficientes visuais.	Representação da informação	2015
D60	Ontologia para um motor de busca semântica para recuperação jurisprudencial no Brasil.	Ontologias	2017
D63	Diretrizes para organização, disseminação e recuperação dos atos normativos da Agência Nacional de Mineração.	Gestão documental	2018
D67	Recuperação da informação através de busca comparada em domínio específico, baseado em expressões multipalavras.	Mecanismos de busca	2013
D69	A expansão de queries sobre terminologias biomédicas: uma comparação de artefatos de representação do conhecimento para Recuperação de Informações.	Representação da informação	2020
D76	Identificadores persistentes: aplicabilidade na organização e acesso à informação científica.	Representação da informação	2021
D78	Interfaces entre a Ciência da Informação e inteligência artificial: um estudo de caso do Sistema <i>SmartJud</i> do processo judicial eletrônico do Tribunal de Justiça do Estado da Bahia.	Inteligência artificial	2021
D83	Encontrabilidade da informação: contributo para uma conceituação no campo da Ciência da Informação.	Epistemologia	2013
D85	Reconhecimento de emoções em músicas por meio de sistemas neuro-fuzzy.	Modelo de recuperação	2019
D86	SRDigital: proposta de um modelo baseado na linguagem natural e controlada como instrumentos de apoio ao agente computacional do processo de referência.	Modelo de recuperação	2012
D88	Comportamento informacional de profissionais no domínio da saúde: um estudo junto ao Hospital de Reabilitação de Anomalias Craniofaciais da Universidade de São Paulo.	Competência em informação	2015
D93	Recuperação e classificação de informações provenientes da <i>web</i> e de redes sociais.	<i>Web</i> e redes sociais	2013
D94	Uma proposta de classificação facetada para organização e recuperação da informação do acervo técnico da Força Aérea Brasileira.	Representação temática	2017

Quadro 3 – Produção acadêmico-científica analisada.

3 de 3

ID	Título	Classificação (Temáticas)	Ano de defesa
D115	Ciência da informação em contextos de e-science: bibliotecários de dados em tempos de Data Science.	<i>Data Science</i>	2017
D131	Extração de dados <i>web</i> como suporte na elaboração de indicadores do turismo de Minas Gerais: uma iniciativa em <i>Big Data</i> .	Big data	2017
D137	Wikis semânticos: da <i>Web</i> para a <i>Web Semântica</i> .	<i>Web</i> semântica	2014
D138	Motor de busca aberto como estratégia de indexação e mineração de dados.	Mineração de dados	2015
D143	Contribuição da Ciência da Informação para os estudos de prospecção de informação científica e tecnológica.	Epistemologia	2019
D144	As coleções de plantas em herbários: a organização e representação da informação sob aspectos históricos e parâmetros metodológicos.	Representação da informação	2017
D149	Comportamento de busca de informação dos agentes sociais das entidades da sociedade civil organizada brasileira: a abordagem do ‘horizonte de informação’ no contexto da participação em iniciativas de governo aberto.	Estudo de usuários	2017
D159	Análise de atos normativos: o caso da Agência Nacional de Energia Elétrica – Aneel.	Estudo de usuários	2016
D161	Processamento de linguagem natural para indexação automática semântico-ontológica.	Representação temática	2013
D162	SNIIC: a plataforma digital de cultura brasileira sob a perspectiva das convergências entre o Design e a Ciência da Informação.	Arquitetura da informação	2016
D165	Estudo da padronização visando à interoperabilidade: o caso das organizações de saúde na cidade de Marília - São Paulo.	Interoperabilidade	2015
D177	Tecnologias para aplicação da <i>Web</i> semântica nas unidades de informação.	<i>Web</i> semântica	2015
D184	Ontologias difusas no suporte à mineração de dados: aplicações na Secretaria de Finanças da Prefeitura Municipal de Belo Horizonte.	Ontologias	2012
D194	RISO - TT - extração de expressões temporais em textos.	Mecanismos de busca	2013
D198	A interpretação semântica de textos científicos em português na perspectiva da Ciência da Informação: procedimentos e aplicação à área de Ciências Agrárias.	Mecanismos de busca	2016
D200	Sistema de informação da pós-graduação brasileira: avaliação da Plataforma Sucupira.	Mecanismos de busca	2020
D207	Trilhou: uma aventura gamificada com Maria Livrão no universo da pesquisa escolar no ensino fundamental.	Mecanismos de busca	2021
D214	Um modelo de interface extensível para sistemas de registro eletrônico de saúde baseados na norma ISO 13606.	Modelo de recuperação	2016
D216	Implementando o prontuário eletrônico <i>OpenEHR</i> em CMSs: uma aproximação.	Interoperabilidade	2014
D229	Método ágil de integração semântica de dados científicos baseado em ontologias.	Interoperabilidade	2020
D235	Um estudo acerca dos recursos audiovisuais no contexto do <i>Linked Data</i> .	Linked data	2016
D240	Estudo sobre a incorporação de dados em bibliotecas digitais de teses e dissertações utilizando <i>linked data</i> .	Linked data	2017
D241	Recuperação da informação científica e tecnológica à luz da prática clínica baseada em evidência: proposta de desenvolvimento de modelo para aplicação em telessaúde.	Modelo de recuperação	2012

Fonte: Elaborada pela autora. Belo Horizonte (2023)

Todos os documentos foram disponibilizados no repositório desta pesquisa e são indicados por um identificador “ID”, número sequencial criado para gestão dos dados desta pesquisa. Destaca-se que, para três trabalhos, o acesso não era aberto, e foi necessário realizar um cadastro na plataforma para ter acesso ao documento completo. Um trabalho estava com o *link* incorreto, no entanto, em uma busca simples na *internet*, foi possível localizar o trabalho completo.

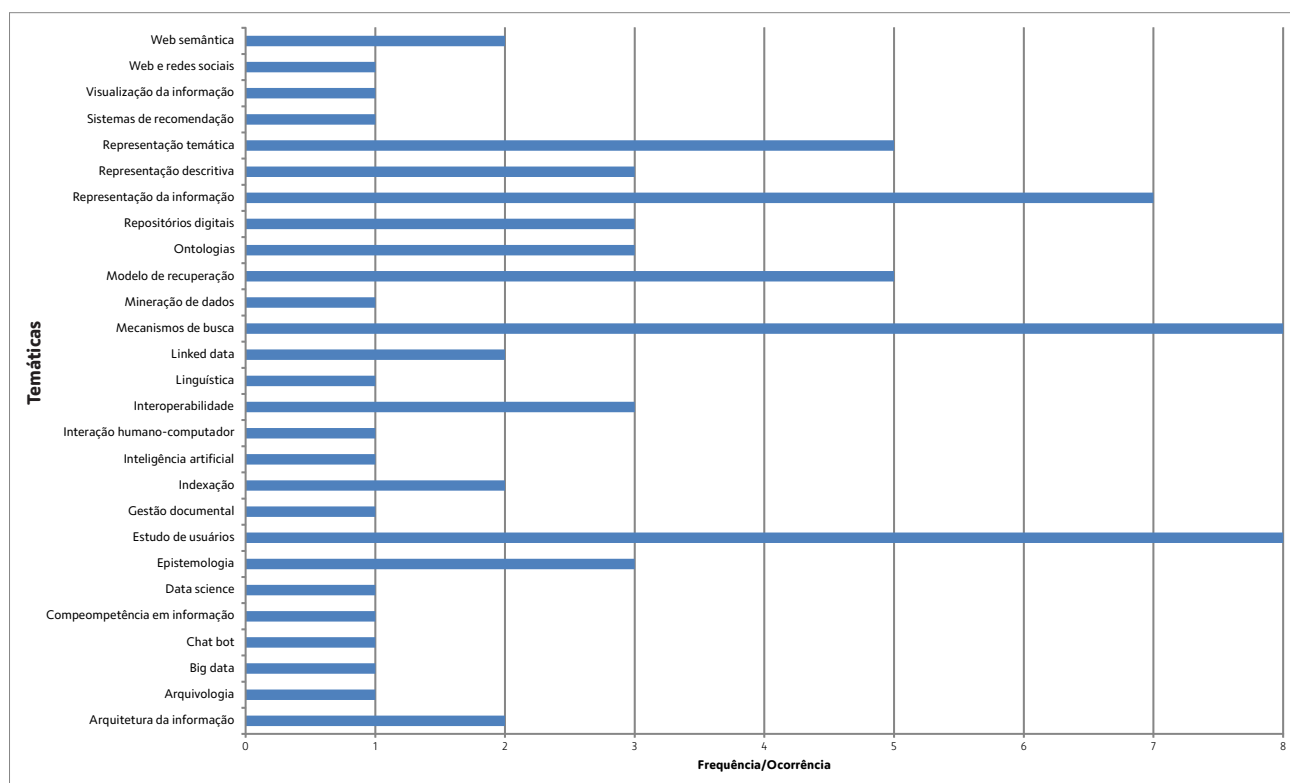


Figura 1 – Produção acadêmico-científica da Ciência da Informação sobre Recuperação de Informação (Temáticas).

Fonte: Elaborado pela autora. Belo Horizonte (2023).

Resultados

Para a amostra selecionada, 69 trabalhos, foram identificadas 27 temáticas: arquitetura da informação; arquivologia; *Big Data*; *chatbot*; competência em informação; *data science*; epistemologia; estudo de usuários; gestão documental; indexação; inteligência artificial; interação humano-computador; interoperabilidade; linguística; *Linked Data*; mecanismos de busca; mineração de dados; modelo de recuperação; ontologias; repositórios digitais; representação da informação; representação descritiva; representação temática; sistemas de recomendação; visualização da informação; *Web* e redes sociais e *Web* semântica. A relação dos trabalhos é apresentada no Quadro 3, e sua frequência é apresentada na Figura 1.

As temáticas identificadas nos trabalhos analisados são apresentadas na Figura 1 e exemplificam a característica multidisciplinar da recuperação de informação, que está relacionada com diversas temáticas, na Ciência da Informação ou em outras áreas. O maior número de trabalhos foi identificado para as temáticas: Estudo de usuários (8), Mecanismos de busca (8), Representação da informação (7), Modelo de Recuperação (5) e Representação temática (5).

O Quadro 4 apresenta a distribuição das temáticas ao longo dos anos analisados. Observou-se que as temáticas com o maior número de trabalhos, como estudos de usuários e mecanismos de busca, estão bem distribuídas entre 2012 e 2021; todavia, a temática representação da informação desponta a partir de 2015. Temáticas mais recentes como *chatbot* e inteligência artificial são observadas a partir de 2020.

A Figura 2 apresenta a vinculação dos trabalhos analisados. A Universidade Federal de Minas Gerais (UFMG) foi a instituição com mais trabalhos (14), seguida da Universidade Estadual Paulista (UNESP) e Universidade de São Paulo (USP), com 9 trabalhos, e a Universidade Federal de Santa Catarina (UFSC) com 7 trabalhos.

Quadro 4 – Distribuição das temáticas ao longo dos anos.

Classificação	Ano de defesa
Arquitetura da informação	2014, 2016
Arquivologia	2019
Big data	2017
Chatbot	2020
Competência em informação	2015
Data science	2017
Epistemologia	2013, 2015, 2019
Estudo de usuários	2012, 2014, 2015, 2016(3), 2017, 2019
Gestão documental	2018
Indexação	2019 (2)
Inteligência artificial	2021
Interação humano-computador	2013
Interoperabilidade	2014, 2015, 2020
Linguística	2015
Linked data	2016, 2017
Mecanismos de busca	2013 (2), 2014, 2015(2), 2016, 2020, 2021
Mineração de dados	2015
Modelo de recuperação	2012(2), 2016(2), 2019
Ontologias	2012, 2015, 2017
Repositórios digitais	2012, 2013, 2015
Representação da informação	2015(2), 2016, 2017, 2018, 2020, 2021
Representação descritiva	2012, 2015, 2017
Representação temática	2013(3), 2014, 2017
Sistemas de recomendação	2017
Visualização da informação	2019
Web e redes sociais	2013
Web semântica	2014, 2015

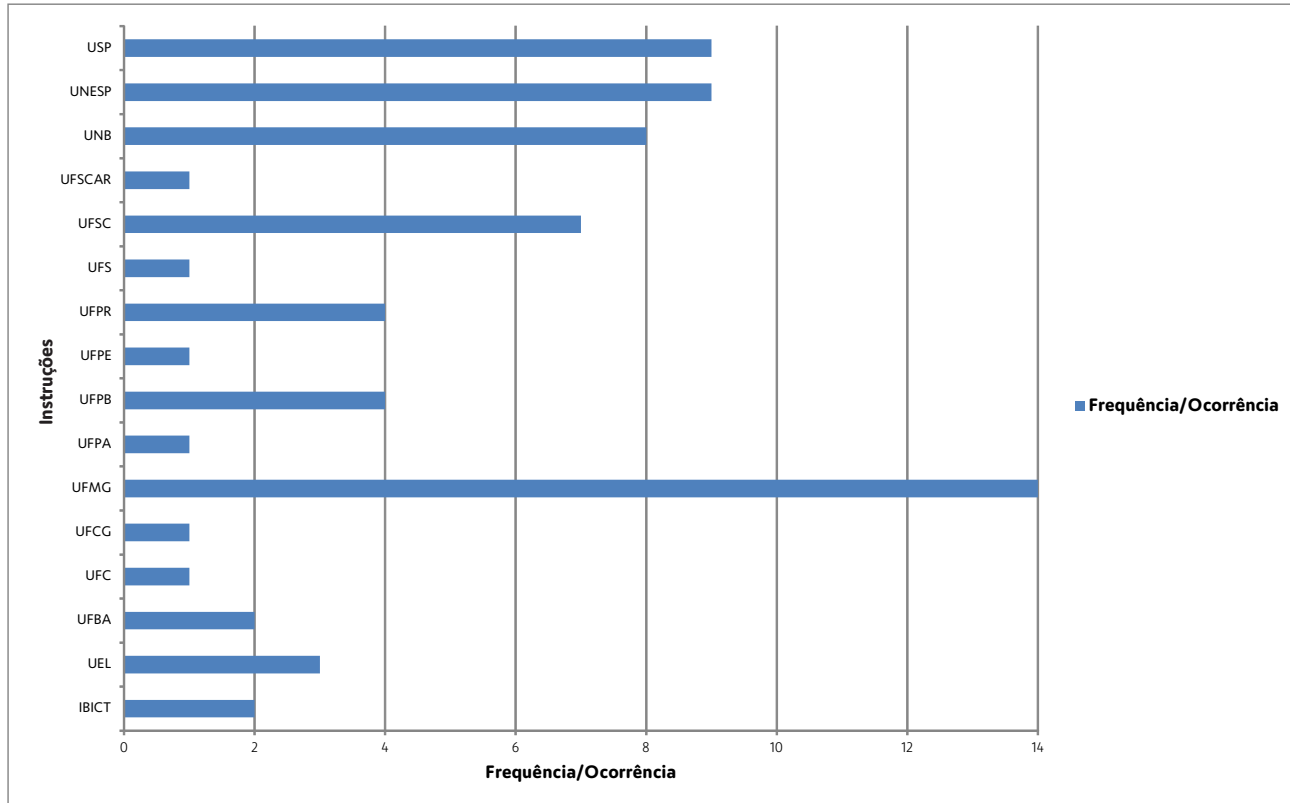


Figura 2- Produção acadêmico-científica da Ciência da Informação sobre Recuperação de Informação (Vinculação).

Fonte: Elaborado pela autora. Belo Horizonte (2023).

Discussão

Na temática estudo de usuários, as publicações estão distribuídas entre 2012 e 2019 e apresentam a interação, as estratégias e o comportamento de busca dos usuários em casos de uso específicos (D17, D20, D26, D34, D42, D49, D149, D159). O trabalho D34, por exemplo, investiga a recuperação de informação em uma base de dados da saúde, por médicos residentes, analisando a interface com uma avaliação heurística que inclui elementos de arquitetura e recuperação de informação. Apesar de a recuperação de informação ter origem na Ciência da Computação (Baeza-Yates; Ribeiro-Neto, 2013), seus objetivos e resultados estão na Ciência da Informação, por promover o acesso e atender as necessidades de informação do usuário. Essa característica pode explicar o elevado número de trabalhos nesta temática.

Na temática mecanismos de busca, o trabalho D11 apresenta um estudo de caso avaliando interfaces de busca de jornais *online*, os trabalhos D41, D47 e D200 apresentam estudos de caso em bases de dados, portal e plataforma, respectivamente, envolvendo a perspectiva do usuário e a definição de critérios para busca e recuperação. A perspectiva da recuperação de informação, em outros recursos, como os jornais *online*, desponta em 2015, evoluindo para recuperação de informação em plataformas, em 2020. Novos contextos também são identificados: o trabalho D198 analisa a busca semântica por meio da extração de relacionamentos semânticos do tipo “causa e efeito” em artigos científicos da área de Ciências Agrárias, enquanto os trabalhos D67 e D194 analisavam a recuperação de informação utilizando padrões de busca comparada em domínio específico e padrões temporais, respectivamente. Por fim, o trabalho D207 traz uma pesquisa aplicada, cujo produto final foi um *game*, com o objetivo de facilitar o entendimento das necessidades de busca e uso da informação em uma biblioteca escolar. Observou-se que o ambiente tradicional das bibliotecas ganha novas possibilidades e deixa de ser dominante.

Os trabalhos que abordam as temáticas de representação da informação (D10, D21, D46, D57, D69, D76 e D144), representação descritiva (D4, D36 e D38) e representação temática (D39, D43, D94, D161) abordam métodos e técnicas de representação da informação visando garantir a recuperação de informação, ou como consequência para a recuperação e acesso à informação, identificando intercorrências que interferem e impactam na recuperação de informações. Nestes trabalhos, discute-se: catalogação e metadados, vocabulário controlado, tesouros, indexação, linguagens documentárias dentre outras técnicas e métodos. No trabalho D4, Ferreira (2015), apresenta no seu referencial teórico o processo de recuperação de informação difundindo os modelos clássicos (booleano, vetorial, probabilístico) e as técnicas de recuperação de informação como *relevance feedback*, filtragem, recomendação e *clustering* ou agrupamento. No entanto, seu objetivo principal está relacionado à representação musical por meio de metadados ID3.

Conforme abordado por Mooers (1951), a recuperação de informação estabeleceu-se como um campo do conhecimento científico com o objetivo de solucionar três problemas enunciados pela explosão da informação: (1) representação da informação; (2) especificação da busca da informação e (3) criação de mecanismo para recuperação. Nesse mesmo sentido, e baseado nos estudos de Robredo e HjØrland, define-se a recuperação de informação como:

[...] um campo que abrange diversos domínios, desde a Ciência da Informação até Ciência da Computação, *possui ferramentas de organização e recuperação da informação e conhecimento, como: classificação, tesouros, taxonomia e ontologias*, nos quais é possível estabelecer diferentes estudos teóricos e múltiplas visões do desdobramento de suas aplicações tidas como fundamentais na RI (Pontes Junior; Carvalho; Azevedo, 2013, p. 6, grifo nosso).

Estas definições justificam, de certo modo, a quantidade de trabalhos na temática de representação e relação existente com a recuperação de informação. É importante destacar também que tais trabalhos costumam apresentar em seu referencial teórico os fundamentos da recuperação de informação e da representação.

Com relação à temática Modelo de Recuperação, foram identificados cinco trabalhos entre 2012 e 2019. Dentre esses, o trabalho D241 propõe um modelo de recuperação de informação científica, no entanto é uma tese na área de saúde pública e não apresenta os fundamentos teóricos da área de recuperação de informação, apesar de fazer uma análise de bases de dados da área da saúde. Os outros trabalhos (D7, D85, D86 e D214) propõem o desenvolvimento ou a aplicação de modelos, fundamentado nos modelos clássicos de recuperação de informação e novos elementos como: redes complexas e dados abertos ligados no trabalho D7, o contexto musical por meio da lógica *fuzzy* no trabalho D85, técnicas de processamento de linguagem natural no trabalho D86 e a Norma 13606, no trabalho D214, que define padrões de comunicação para o registro eletrônico em saúde.

A temática interoperabilidade é abordada no contexto da saúde por meio de padrões para comunicação de registros de saúde (D165, D216) e na integração semântica baseada em ontologias de domínio (D229). A temática ontologia também é apresentada nos trabalhos (D25, D60, D184). O primeiro trabalho investiga as tecnologias da *Web* semântica e as técnicas de processamento de linguagem natural para elaboração semiautomática de uma base léxico-ontológica, no domínio de risco financeiro, para ser implementada ao modelo de recuperação de informação. O segundo trabalho tem como o propósito desenvolver uma ontologia, para construção de um motor de busca no contexto jurisprudencial. O terceiro trabalho aborda a mineração de dados, aplicada na Secretaria de Finanças da Prefeitura de Belo Horizonte. As três soluções apresentaram resultados positivos na recuperação de informações, porém em um contexto muito específico. Em todo caso, espera-se que o método possa ser adaptado para outros cenários.

Os trabalhos classificados com a temática *Linked Data* (D235, 240) abordam a ligação de dados, na estrutura de padrões de descrição e representação de recursos, e no tratamento da informação por meio da integração com ontologias e o Projeto DBPedia, conhecido pela diversidade de conjuntos de dados abertos.

Alguns trabalhos destacam-se por abordar fenômenos contemporâneos como *Big Data* (D131), *Data Science* (D115), visualização da informação (D13) e novos recursos como o *chatbot* (D1). No trabalho D13, intitulado: “Extroversão e descoberta: visualização aplicada em sistemas de recuperação da informação”, Barbosa (2019) apresenta uma discussão importante na área associando a Ciência da Informação à visualização da informação. Para tanto, o trabalho apresenta o contexto atual descrevendo e caracterizando o fenômeno *Big Data* em paralelo com o contexto histórico da Ciência da Informação e da recuperação da informação. Em um segundo momento, o trabalho apresenta os conceitos de visualização da informação, algumas técnicas empregadas na Ciência da Informação e outros contextos e, por fim, um estudo de caso em uma empresa pública.

O trabalho D115 intitulado: “Ciência da informação em contextos de *e-science*: bibliotecários de dados em tempos de *Data Science*” apresenta conceitos recentes e discute o papel do bibliotecário, principalmente em repositórios de dados de pesquisa. O estudo destaca algumas técnicas de manipulação de dados como: classificação, *clustering*, regressão linear, simulação, análise de conteúdo e mecanismos de recomendação, assim como alguns tipos de análise: descritiva, diagnóstico, preditiva e prescritiva. A recuperação da informação é destacada em uma seção que apresenta conceitos sobre a recuperação na internet, com destaque para técnicas de *web scraping*, coleta de dados e o *harvesting* de metadados, uma coleta feita por protocolo,

através de requisições, permitindo a interoperabilidade entre sistemas. Ferramentas recentes para manipulação de dados também são apresentadas. Por fim, a pesquisa apresenta o cenário internacional dos repositórios de dados de pesquisa em geociências.

No trabalho D1 intitulado “Batendo um papo com a informação: o uso dos *chatbots* para a recuperação da informação e a contribuição da Ciência da Informação nesse processo”, Silva (2020) faz uma abordagem diferente da recuperação da informação, apresentando o *chatbot*, sua arquitetura, estudos de caso e a aplicação prática com técnicas da Ciência da Informação. As entidades e as intenções, elementos que compõem a estrutura do *chatbot*, são apresentadas sob a óptica da indexação, tesouros, taxonomias e ontologias, a fim de padronizar as perguntas e as respostas e gerenciar fluxos de diálogos. Trata-se de uma associação importante que destaca a aplicabilidade do tratamento temático da informação em uma ferramenta pautada na linguagem natural e que pode ser utilizada em diversos ambientes informacionais. Esse trabalho é muito importante para destacar que a Ciência da Informação está presente em diversas áreas, ampliando a atuação do profissional da informação. No trabalho D78 intitulado: “Interfaces entre a Ciência da Informação e inteligência artificial: um estudo de caso do Sistema SmartJud do processo judicial eletrônico do Tribunal de Justiça do Estado da Bahia”, Fontoura (2021) também apresenta um estudo de caso de um *chatbot* integrando a recuperação da informação com técnicas de inteligência artificial no contexto da Ciência da Informação.

A *Web* e a *Web* semântica são abordadas nos trabalhos (D93, D137 e D177). Um estudo exploratório e metodológico sobre o problema de recuperação e classificação na internet é realizado no trabalho D93, que indica o uso de gráficos sociais e a mineração de opiniões para serem utilizados juntamente com técnicas de ranqueamento de páginas (*pagerank*), melhorando os resultados das buscas. O trabalho D177 descreve tecnologias, recomendações e padrões necessários para a aplicação da *Web* semântica nas unidades de informação, propondo um *framework*. O trabalho D137 aborda os *wikis* semânticos, com enfoque nas atividades de descrição, recuperação e intercâmbio de informações, apoiadas em tecnologias da *Web* semântica.

Dialogando com questões epistemológicas da Ciência da Informação, foram identificados três trabalhos (D8, D83 e D143). No trabalho D8, Nhacuongue (2015) aborda em sua tese a essência da recuperação de informação, que é refletida nas tecnologias pós-modernas:

Por outras palavras, a essência da *RI* incide sobre a necessidade de obter e ordenar documentos relevantes na coleção. Este desafio suscita novas abordagens para cada estágio de desenvolvimento e sofisticação dos *SRRI*, principalmente pela proporção do volume de informações e versatilidade da tecnologia de produção, processamento, armazenamento e disseminação (Nhacuongue, 2015, p. 119, grifo nosso).

Dessa forma, percebe-se que as tecnologias evoluem, mas a essência e as necessidades permanecem. A recuperação de informação em sistemas tem evoluído dentro dos conceitos da interoperabilidade, que vislumbra, por exemplo, novas interfaces de programação de aplicativos (Application Programming Interface, API), uma tendência já implementada em diferentes sistemas de informação para troca de informações (Silva, 2022). Porém, as necessidades de busca do usuário precisam ser atendidas.

Observou-se, por meio da classificação realizada e da leitura dos trabalhos, que muitas pesquisas estão relacionadas à organização e à representação da informação, áreas totalmente interligadas com a recuperação da informação e necessárias para que a recuperação seja efetiva. Na temática de indexação, o trabalho D19 apresenta a relevância da política de indexação no processo

de recuperação de informação por meio de uma revisão sistemática de literatura. Nesse mesmo sentido, o trabalho D30 faz uma avaliação da indexação por meio da exaustividade e especificidade pela revocação e precisão na recuperação de informação. Fujita (2012, *apud* Garcia, 2019, p. 13) destaca que a indexação e a recuperação da informação apresentam laços estreitos como uma relação de causa e efeito, já que a indexação “[...] possui características que causam efeitos na recuperação da informação”. Destaca-se que o trabalho D138, intitulado: “Motor de busca aberto como estratégia de indexação e mineração de dados”, faz um caminho inverso, partindo da recuperação para orientar a representação, ao analisar um motor de busca como estratégia para indexação e mineração de dados. Este trabalho aponta a forte ligação entre a organização, a representação e o armazenamento, elementos que compõem um sistema de recuperação de informação.

Os trabalhos com a temática denominada repositórios digitais (D2, D6, D14) destacam problemas de representação e visam contribuir com a recuperação em acervos bibliográficos, assim como as temáticas de arquivologia (D3) e gestão documental (D63). Temáticas como competência em informação (D88), linguística (D54), arquitetura da informação (D23, D162), interação humano-computador (D15) e sistemas de recomendação (D24) apresentam uma intersecção com a área de recuperação de informação e, por isso, estão associadas aos trabalhos selecionados. As principais contribuições estão focadas na interação com o usuário.

Após a análise de todos os 69 trabalhos selecionados, é importante salientar que somente um trabalho não era da área de Ciência da Informação, sendo excluído das análises seguintes. Dessa forma, a amostra final foi composta por 68 trabalhos, sendo 21 teses e 47 dissertações. Com relação às temáticas identificadas e ao ano da defesa dos respectivos trabalhos, observa-se que temáticas como *Big Data*, *Data Science*, *Linked Data*, inteligência artificial e *chatbot* são identificadas a partir de 2016. Já as temáticas que envolvem a representação da informação são distribuídas quase que uniformemente nos últimos 10 anos. As instituições com mais trabalhos sobre recuperação de informação (UFMG, UNESP, USP, UFSC) estão localizadas na região sudeste, com exceção da UFSC na região sul.

Considerações Finais

A recuperação de informação, apesar de implementada há milênios nas bibliotecas, é uma área recente que concentra estudos da Ciência da Computação e da Ciência da Informação. O termo, criado a partir dos anos 1950, intensificou os estudos na área e tem acompanhado as evoluções tecnológicas. Com sua característica multidisciplinar, a recuperação de informação dialoga com diferentes temáticas e tem se destacado na solução de problemas e fenômenos atuais, envidando técnicas e métodos de diferentes áreas.

No contexto da Ciência da Informação, a recuperação de informação tem forte ligação com a organização e a representação da informação e compõe uma relação quase que indissociável. Ainda sobre essa relação, é importante destacar que ela é fortemente refletida na produção acadêmica da área, ao longo dos 10 anos analisados, e traz certa “identidade”, quando analisada no âmbito da Ciência da Informação, não se restringindo ao problema de recuperação de informação, mas também a uma etapa anterior, que está associada à organização e ao tratamento da informação.

A evolução dos recursos tecnológicos e seu impacto na recuperação de informação são percebidos com a inserção de pesquisas envolvendo, principalmente, técnicas de inteligência artificial e algoritmos. Além disso, percebe-se também que há uma tendência de pesquisas envolvendo contextos diferenciados (saúde pública, ambientes corporativos, poder judiciário) e não somente bibliotecas, arquivos e museus. A identificação de 27 temáticas relacionadas à

recuperação de informação demonstra a multidisciplinaridade da área. A região sudeste do país se destaca na quantidade de trabalhos sobre recuperação de informação, aproximação que pode estar relacionada ao desenvolvimento econômico da região.

Em resumo, o objetivo proposto nesta pesquisa foi alcançado ao apresentar um panorama da produção acadêmico-científica brasileira sobre a área de recuperação de informação, na Ciência da Informação, entre 2012 e 2021. Apesar das limitações existentes, o estudo apresentou uma fotografia do contexto atual, tornando-se um aporte para novos pesquisadores e uma oportunidade para a comunidade acadêmico-científica refletir sobre as contribuições existentes, instigando novas reflexões sobre a recuperação de Informação na Ciência da Informação.

Referências

- Araújo, C. A. A. A Ciência da Informação como Ciência Social. *Ciência da informação*, v. 32, n. 3, p. 21-27, 2003. Disponível em: <http://www.scielo.br/pdf/ci/v32n3/19020>. Acesso em: 20 dez. 2022.
- Baeza-Yates, R.; Ribeiro-Neto, B. *Recuperação de informação: conceitos e tecnologia das máquinas de busca*. 2. ed. Porto Alegre: Bookman, 2013.
- Barbosa, E. C. *Extroversão e descoberta: visualização aplicada em sistemas de recuperação da informação*. 2019. 168 f. Dissertação (Mestrado em Ciência da Informação) – Universidade de São Paulo, São Paulo, 2019. Disponível em: <http://www.teses.usp.br/teses/disponiveis/27/27151/tde-22012020-165026/>. Acesso em: 28 jan. 2023.
- Berners-Lee, T.; Hendler, J.; Lassila, O. The Semantic Web. *Scientific American*, p. 34-43, 2001. Disponível em: <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>. Acesso em: 28 jan. 2023.
- Cesarino, M. A. N. Sistemas de recuperação da informação. *Revista da Escola de Biblioteconomia da UFMG*, v. 14, n. 2, 1985. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/76506>. Acesso em: 28 jan. 2023.
- Ferneda, E. *Recuperação de informação: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação*. 2003. Tese (Doutorado em Ciência da Informação e Documentação) – Universidade de São Paulo, São Paulo, 2003. Disponível em: <https://www.teses.usp.br/teses/disponiveis/27/27143/tde-15032004-130230/pt-br.php>. Acesso em: 28 jan. 2023.
- Ferneda, E. *Introdução aos modelos computacionais de recuperação de informação*. Rio de Janeiro: Ciência Moderna, 2012.
- Ferreira, J. B. *Recuperação de informação de música e dados ID3: possíveis aplicações*. 2015. 62 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, São Paulo, 2015. Disponível em: https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ferreira_ju_me_mar.pdf. Acesso em: 28 jan. 2023.
- Fontoura, R. V. *Interfaces entre a Ciência da Informação e inteligência artificial: um estudo de caso do Sistema SmartJud do processo judicial eletrônico do Tribunal de Justiça do Estado da Bahia*. 2021. 105 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal da Bahia, Bahia, 2021. Disponível em: <https://repositorio.ufba.br/handle/ri/35052>. Acesso em: 28 jan. 2023.
- Garcia, V. C. *Avaliação da indexação a partir dos elementos de exaustividade e especificidade pela revocação e precisão na recuperação da informação: sob a ótica dos usuários no sistema de bibliotecas da UFPA*. 2019. 185 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal do Pará, Belém, 2019. Disponível em: <http://repositorio.ufpa.br/jspui/handle/2011/11435>. Acesso em: 16 jan. 2023.
- Instituto Brasileiro de Informação em Ciência e Tecnologia. *Biblioteca Digital Brasileira de Teses e Dissertações*. Brasília: IBICT, 2023. Disponível em: <https://btd.ibict.br/vufind/>. Acesso em: 19 jan. 2023.
- Mendes, K.D.S.; Silveira, R.C.C.P.; Galvão, C.M. Revisão integrativa: método de pesquisa para a incorporação de evidências na saúde e na enfermagem. *Texto & Contexto: Enfermagem*, v. 17, n. 4, p. 758-764, 2008. Disponível em: <https://www.scielo.br/j/tce/a/XzFkq6tjWs4wHNqNjKlLkXQ/>. Acesso em: 23 jan. 2023.
- Mooers, C. Zatocoding applied to mechanical organization of knowledge. *American Documentation*, v. 2, n. 1, p. 20-32, 1951.

Nhacuongue, J. A. *O campo da Ciência da Informação: contribuições, desafios e perspectivas da mineração de dados para o conhecimento pós-moderno*. 2015. 194f. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2015. Disponível em: <https://repositorio.unesp.br/handle/11449/126578?show=full>. Acesso em: 28 jan. 2023.

Oliveira, H. S.; Cavalcante, L. E. Ex-Libris: uma revisão integrativa. *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, v. 20, p. e022022, 2022. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8670945>. Acesso em: 16 jan. 2023.

Paes, D. M. B.; Tabosa, H. R. Biblioteca digital de teses e dissertações: reflexões sobre representação da informação com vistas à recuperação da informação. *Revista ACB: Biblioteconomia em Santa Catarina*, v. 20, n. 2, p. 225-239, 2015. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/73593>. Acesso em: 24 jan. 2023.

Pontes Junior, J.; Carvalho, R. A.; Azevedo, A. W. Da recuperação da informação à recuperação do conhecimento: reflexões e propostas. *Perspectivas em Ciência da Informação*, v. 18, n. 4, p. 2-17, 2013. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/35875>. Acesso em: 28 jan. 2023.

Saracevic, T. Information Science. *Journal of the American Society for Information Science*, v. 50, n. 12, p. 1051-1063, 1999.

Saracevic, T. Ciência da Informação: origens, evolução e relações. *Perspectivas em Ciência da Informação*, v. 1, n. 1, p. 41-62, 1996. Disponível em: <https://periodicos.ufmg.br/index.php/pci/article/view/22308>. Acesso em: 28 jan. 2023.

Silva, R. J. *Batendo um papo com a informação: o uso dos chatbots para a recuperação da informação e a contribuição da Ciência da Informação nesse processo*. 2020. 125 f. Dissertação (Mestrado em Organização, Mediação e Circulação da Informação) – Universidade de São Paulo, São Paulo, 2020. Disponível em: <https://www.teses.usp.br/teses/disponiveis/27/27163/tde-10032021-013140/publico/RafaelJoadaSilvaVC.pdf>. Acesso em: 28 jan. 2023.

Silva, P. N. Recuperação de dados em APIs: uma experiência prática no Twitter. In: Encontro Nacional de Pesquisa em Ciência da Informação, 22., 2022, Porto Alegre. *Anais [...]*. Porto Alegre: Ancib, 2022. Disponível em: <https://ancib.org/enancib/index.php/enancib/xxiiencib/paper/viewFile/1007/728>. Acesso em: 27 abr. 2023.

Wersig, G.; Neveling, U. The phenomena of interest to Information Science. *The informationscientist*, v. 9, n. 4, p. 127-140, 1975. Disponível em: <https://sigir.org/files/museum/pub-13/18.pdf>. Acesso em: 28 jan. 2023.

Whittemore, R.; Knafelz, K. A. The integrative review: updated methodology. *Journal of Advanced Nursing*, v. 52, n. 5, p. 546-553, 2005.